DOCUMENT RESUME

ED 440 996		TM 030 794
AUTHOR	Headrick, Todd C.; Sawilowsky, Shlomo S.	
TITLE	Properties of the Rank Transformation in Factorial Analysis of Covariance.	
PUB DATE	2000-04-00	
NOTE	31p.; Paper presented at the Annual Meeting of the American Educational Research Association (New Orleans, LA, April 24-28, 2000).	
PUB TYPE	Numerical/Quantitative Data (110) -- Reports - Evaluative (142) -- Speeches/Meeting Papers (150)	
EDRS PRICE	MF01/PC02 Plus Postage.	
DESCRIPTORS	*Analysis of Covariance; *Factor Analysis; Factor Structure; Nonparametric Statistics	
IDENTIFIERS	*Rank Order Transformation; Type I Errors	

ABSTRACT

Real world data often fail to meet the underlying assumption of population normality. The Rank Transformation (RT) procedure has been recommended as an alternative to the parametric factorial analysis of Covariance (ANCOVA). The purpose of this study was to compare the Type I error and power properties of the RT ANCOVA to the parametric procedures in the context of a completely randomized balanced 3 x 4 factorial layout with one covariate. This study was concerned with tests of homogeneity of regression coefficients and interaction under conditional (non)normality. Both procedures displayed erratic Type I error rates for the test of homogeneity of regression coefficients under conditional nonnormality. With all parametric assumptions valid, the simulation results demonstrate that the RT ANCOVA failed as a test for either homogeneity of regression coefficients or interaction due to severe Type I error inflation. The error inflation was most severe when departures from conditional normality were extreme. Also associated with the RT procedure was a loss of power. It is recommended that the RT procedure not be used as an alternative to factorial ANCOVA despite its encouragement from publishers of statistical software packages. (Contains 19 tables and 38 references.) (Author/SLD)

PROPERTIES OF THE RANK TRANSFORMATION IN FACTORIAL ANALYSIS OF COVARIANCE

Todd C. Headrick
Measurement & Statistics
223D Wham Building
Southern Illinois University
Carbondale, IL 62981-4618

Shlomo S. Sawilowsky
Evaluation & Research
351 College of Education
Wayne State University
Detroit, MI 48202

Key Words: Type I error; power; nonparametric; designed experiments

ABSTRACT

Real world data often fail to meet the underlying assumption of population normality. The Rank Transformation (RT) procedure has been recommended as an alternative to the parametric factorial analysis of covariance (ANCOVA). The purpose of this study was to compare the Type I error and power properties of the RT ANCOVA to the parametric procedure in the context of a completely randomized balanced 3×4 factorial layout with one covariate. This study was concerned with tests of homogeneity of regression coefficients and interaction under conditional (non)normality. Both procedures displayed erratic Type I error rates for the test of homogeneity of regression coefficients under conditional nonnormality. With all parametric assumptions valid, the simulation results demonstrated that the RT ANCOVA failed as a test for either homogeneity of regression coefficients or interaction due to severe Type I error inflation. The error inflation was most severe when departures from conditional normality were extreme. Also associated with the RT procedure was a loss of power. It is recommended that the RT procedure not be used as an alternative to factorial ANCOVA despite its encouragement from SAS, IMSL, and other respected sources.

1. INTRODUCTION

The Rank Transformation (RT) procedure is appealing due to its simplicity and ease of execution. The steps for testing hypotheses are: (1) replace the raw scores with their respective rank order, (2) conduct the classical normal theory tests on the ranks, and (3) refer the test statistic to the usual tables of percentage points. Most statistics software packages contain the parametric tests and the ranking or sorting routines necessary to easily conduct the RT procedure.

Conover and Iman (1982) promoted the use of the RT as an alternative to the parametric analysis of covariance. This was an elaboration stemming from their overview of the RT (Conover & Iman, 1981). The extension of the RT to ANCOVA was based on Monte Carlo results from factorial ANOVA (Iman, 1974; Conover & Iman 1976) and regression (Iman & Conover, 1979) designs.

Conover and Iman (1982) presented a general linear model using dummy coding to consider both total-group and within-group regression slopes. Thus, the RT ANCOVA was also suggested for the purpose of testing the assumption of homogeneity of regression coefficients. Most other competing nonparametric procedures (e.g., Hettmansperger, 1984; Puri & Sen, 1969a; Quade, 1967) consider only the total-group regression slope. As such, the condition of equal slopes must be assumed to hold in order to safely use the competing procedures.

Conover and Iman (1982) asserted, "The rank transform procedure can be extended beyond the one-way ANCOVA by further use of dummy-variables" (p. 723). Subsequently, Conover and Iman (1982) provided an example of how to extend the RT procedure to factorial ANCOVA using a 2×3 layout with one covariate and with three observations per cell. They described tests for: (a) the full model, (b) homogeneity of regression slopes, (c) interaction, and (d) main effects.

Empirical research by Harwell and Serlin (1988), Olejnik and Algina (1984, 1987), Seaman, Algina, & Olejnik (1985), and Stephenson and Jacobson (1988) found conditions favorable for the one-way RT ANCOVA. Extension to the factorial case was also postulated. For example, Olejnik and Algina (1985), stated: "The rank transformation can be used in factorial designs and with multiple covariates" (p. 62). Similarly, Harwell and Serlin (1988) claimed: "Under the rank transformation principle of Conover and Iman (1981), rank ANCOVA could be extended to the multiple covariate/factorial ANCOVA case, although its asymptotic distributional properties would be unknown" (p. 271).

More recently, Deshon and Alexander (1996) also suggested the use of the RT for the test of homogeneity of regression slopes. Recent suggestions promoting the use of the RT in factorial designs have also been made (Choi, 1998; Potvin & Roff, 1993; Regeth & Stine, 1998). For example, Regeth and Stine (1998) submitted, "for two-way designs (involving an interaction), the ANOVA test can be run, using the rank orderings of data points rather than the actual scores" (p. 708). Further, some recent examples of applications of the RT in complex designs include multiple regression (Angermeier & Winston, 1998) and factorial ANOVA (Augner, Provenza, & Villalba, 1998). It should be noted that these suggestions and applications of the Rank Transform to the general linear model have been made despite studies that have demonstrated numerous limitations of the RT in complex designs (e.g., Akritas, 1990; Brunner & Neumann, 1986; Sawilowsky, Blair, & Higgins, 1989; Thompson, 1991).

2. PURPOSE OF THE STUDY

The purpose of this study is to investigate the Rank Transformation ANCOVA as an alternative to the parametric factorial ANCOVA. It is well known that good nonparametric tests for main effects exist. Thus, this study is concerned with the tests of homogeneity of regression coefficients (slope-treatment interaction) and interaction, either in the presence or absence of main effects with varying degrees of variate and covariate correlation and nonnormality.

3. METHODOLOGY

A completely randomized balanced design with fixed effects and one covariate was used. The structural model representing the design was:

$$Y_{ijk} = \mu + \beta(X_{ijk} - \overline{X}) + \alpha_i + \tau_j + (\alpha\tau)_{ij} + \varepsilon_{ijk}, \qquad (1)$$

($i = 1,..., I$; $j = 1,..., J$; and $k = 1,..., n$), where $I = 3$, $J = 4$, and $n = 3$, 10, and 30.

The levels of variate (Y_{ijk}) and covariate (X_{ijk}) correlation were $\rho = 0$, .3, .6, and .9. By inspection of (1), note that the regression slope coefficient, β, remained constant across groups. Thus, only Type I error was of concern with respect to the test for homogeneity of regression coefficients.

The treatment effect patterns modeled in (1) were as follows:

1. The main effect τ nonnull, the main effect α null, and the interaction $(\alpha\tau)$ null:

   1(a). $\tau_1 = c$;

   1(b). $\tau_1 = \tau_2 = c$; and $\tau_3 = \tau_4 = -c$.

2. The main effects τ and α nonnull, and the interaction $(\alpha\tau)$ null:

   2(a). $\tau_2 = \alpha_1 = c$; and $\tau_3 = \alpha_2 = -c$; and

   2(b). $\tau_3 = \alpha_1 = c$; and $\tau_1 = \tau_2 = \tau_4 = \alpha_3 = -c$.

3. The $(\alpha\tau)$ interaction nonnull, and the main effects τ and α null:

   3(a). $(\alpha\tau)_{11} = (\alpha\tau)_{33} = c$; and $(\alpha\tau)_{13} = (\alpha\tau)_{31} = -c$;

   3(b). $(\alpha\tau)_{11} = (\alpha\tau)_{14} = (\alpha\tau)_{32} = (\alpha\tau)_{33} = c$; and

   $(\alpha\tau)_{12} = (\alpha\tau)_{13} = (\alpha\tau)_{31} = (\alpha\tau)_{34} = -c$.

4. The main effect τ and the $(\alpha\tau)$ interaction nonnull, and the main effect α null:

   4(a). $(\alpha\tau)_{11} = c$; and $(\alpha\tau)_{14} = -c$;

   4(b). $(\alpha\tau)_{11} = (\alpha\tau)_{12} = (\alpha\tau)_{31} = (\alpha\tau)_{32} = c$; and

   $(\alpha\tau)_{13} = (\alpha\tau)_{14} = (\alpha\tau)_{33} = (\alpha\tau)_{34} = -c$.

5. The main effects τ, α, and $(\alpha\tau)$ interaction are nonnull:

5(a). $(\alpha\tau)_{21} = (\alpha\tau)_{24} = c$;

5(b). $(\alpha\tau)_{11} = (\alpha\tau)_{12} = (\alpha\tau)_{32} = (\alpha\tau)_{33} = (\alpha\tau)_{34} = c$; and

$(\alpha\tau)_{13} = (\alpha\tau)_{31} = (\alpha\tau)_{14} = -c$.

The treatment effect sizes (c) ranged from $c = 0.10\sigma$ to $c = 2.00\sigma$, where σ is the standard deviation of the population from which samples were drawn, in increments of 0.10σ. The null case was represented when $c = 0.00$ for all effects.

Five conditional distributions were simulated with zero means, unit variances, and varying amounts of skew (γ_1) and kurtosis (γ_2). A shift parameter was added to model the treatment effects. In all experimental situations, the variate and covariate followed the same distributions. The first conditional distribution selected was a standard normal distribution ($\gamma_1 = 0.0$ and $\gamma_2 = 0.0$). The other four conditional nonnormal distributions simulated were: (a) symmetric and light-tailed ($\gamma_1 = 0.0$ and $\gamma_2 = -1.15132$), (b) symmetric and extremely heavy-tailed ($\gamma_1 = 0.0$ and $\gamma_2 = 25$), asymmetric and moderately heavy-tailed ($\gamma_1 = 1.633$ and $\gamma_2 = 4$), and (d) extremely asymmetric and heavy-tailed ($\gamma_1 = \sqrt{8}$ and $\gamma_2 = 12$).

The steps employed for data generation follow the model developed by Headrick and Sawilowsky (1999) for simulating multivariate nonnormal distributions. This procedure generates correlated nonnormal distributions by combining the Fleishman (1978) power method with a generalization of Theorems 1 and 2 from Knapp and Sowyer (1967).

The Headrick and Sawilowsky (1999) procedure generated the Y_{ijk} and X_{ijk} for the ij-th group in (1) from the use of the following equations:

$$Y_{ijk} = a + bY^{*}_{ijk} + (-a)Y^{*^2}_{ijk} + dY^{*^3}_{ijk} + \delta_{ij}c \text{ and} \tag{2}$$

$$X_{ijk} = a + bX^{*}_{ijk} + (-a)X^{*^2}_{ijk} + dX^{*^3}_{ijk}. \tag{3}$$

The resulting Y_{ijk} and X_{ijk} were distributed with group means of $\delta_{ij}c$ and zero (respectively), unit variances, the desired values of γ_1 and γ_2, and the desired within group correlation (ρ). The value of $\delta_{ij}c$ was the shift parameter added to the *ij*-th group for the treatment effect pattern considered. The coefficients *a*, *b*, and *d* were determined by simultaneously solving Fleishman's Equations 5, 11, 17, 18 (Fleishman, 1978, p. 523) for the desired values of γ_1 and γ_2. The values of Y^*_{ijk} and X^*_{ijk} in (2) and (3) were generated using the following algorithms:

$$Y^*_{ijk} = Z_{ijk}\rho^* + V_{ijk}\sqrt{1-\rho^{*2}} \text{ and} \tag{4}$$

$$X^*_{ijk} = Z_{ijk}\rho^* + W_{ijk}\sqrt{1-\rho^{*2}}, \tag{5}$$

where the Z_{ijk}, V_{ijk}, and W_{ijk} were *N* iid (0,1). The resulting Y^*_{ijk} and X^*_{ijk} were also *N* iid (0,1) and were correlated at the intermediate value ρ^{*^2}. The intermediate correlation, which was different from the desired post-correlation (ρ) except under conditional normality, was determined by solving Equation (7b) for the bivariate case from Headrick and Sawilowsky (1999) for ρ^*.

Values of *a*, *b*, *d*, and ρ^*, were solved using the IMSL subroutine NEQNF (Visual Numerics, 1994, p. 796). These values, along with the post and intermediate correlation values ρ and ρ^{*^2}, are compiled in Table I.

F statistics for main effects, interaction, and homogeneity of regression coefficients were computed on the raw scores and their ranks for the 5 (type of distribution) × 4 (level of variate/covariate correlation) × 3 (sample size) × 21 (treatment effect size) × 10 (treatment pattern) situations. The F statistics were calculated using the sums of squares approach given in Winer, Brown, and Michels (1991) for factorial ANCOVA. Using the F table of percentage points, the proportions of hypotheses rejected for each effect at the .05 and .01 α levels were calculated. Ten thousand repetitions were used for each experiment.

TABLE I

Values of *a*, *b*, *d*, post-correlations (ρ), and intercorrelations (ρ^{*^2}) used in the Headrick and Sawilowky (1999) procedure to simulate the correlated (non)normal distributions.

Distribution	a	b	d	ρ	ρ^{*^2}
1	0.0	1.0	0.0	.00	.00
				.30	.30
				.60	.60
				.90	.90
2	0.0	1.34112	-0.13146	.00	.00
				.30	.330532
				.60	.638209
				.90	.914295
3	0.0	0.25370	0.21380	.00	.00
				.30	.382012
				.60	.689843
				.90	.930689
4	-0.25950	0.88070	0.01621	.00	.00
				.30	.330284
				.60	.631979
				.90	.911219
5	-0.52070	0.61460	0.02007	.00	.00
				.30	.434039
				.60	.712010
				.90	.933780

Note. The distributions are: (1) standard normal ($\mu = 0;\ \sigma = 1;\ \gamma_1 = 0;\ \gamma_2 = 0$);
(2) symmetric and light-tailed ($\mu = 0;\ \sigma = 1;\ \gamma_1 = 0;\ \gamma_2 = -1.15132$);
(3) symmetric and extremely heavy-tailed ($\mu = 0;\ \sigma = 1;\ \gamma_1 = 0;\ \gamma_2 = 25$);
(4) asymmetric and moderately heavy-tailed ($\mu = 0;\ \sigma = 1;\ \gamma_1 = 1.633;\ \gamma_2 = 4$);
(5) extremely asymmetric and heavy-tailed ($\mu = 0;\ \sigma = 1;\ \gamma_1 = \sqrt{8};\ \gamma_2 = 12$).

The computers used to carry out the Monte Carlo were Pentium and Pentium II-based personal computers. All programs were developed using Lahey Fortran 77 version 3.0 (1994), supplemented with various subroutines from RANGEN (Blair, 1987).

4. RESULTS

Type I error and power results are presented in the tables by conditional distribution, sample size, and the treatment effect pattern simulated. The conditional distributions reported are: γ_1 = 0.00 and γ_2 = 0.00; γ_1 = 0.00; γ_2 = −1.1532; γ_1 = $\sqrt{8}$ and γ_2 = 12; and γ_1 = 0.00 and γ_2 = 25. (A complete set of tables is available from the first author.) With respect to Type I error, the effect pattern presented is 2(b): $\tau_3 = \alpha_1 = c$; and $\tau_1 = \tau_2 = \tau_4 = \alpha_3 = -c$, and the sample size presented is n = 30. The effect pattern presented for power analysis is 5(a): $(\alpha\tau)_{21} = (\alpha\tau)_{24} = c$, and the sample sizes presented are n = 3 and n = 10.

The entries in Table II and Table III provide empirical averages of the variate and covariate correlation, skew, and kurtosis with the population parameters used in the simulation. They were obtained by using an averaging procedure described below.

To demonstrate the adequacy of the Headrick and Sawilowsky (1999) procedure, average values of $\rho(\overline{\rho})$, $\gamma_1(\overline{\gamma}_1)$, and $\gamma_2(\overline{\gamma}_2)$ were obtained separately for each sample size and conditional distribution. Values of $\overline{\rho}$ are reported in Table II and values of $\overline{\gamma}_1$ and $\overline{\gamma}_2$ are reported in Table III.

For each repetition, separate values of ρ_{ij}, $\gamma_{1_{ij}}$, and $\gamma_{2_{ij}}$ for each of the IJ groups were computed for the variate and covariate. The average value of $\rho_{ij}(\overline{\rho}_{..})$ was obtained by averaging the ρ_{ij} across the IJ groups. Average values of $\gamma_{1_{ij}}(\overline{\gamma}_{1..})$ and $\gamma_{2_{ij}}(\overline{\gamma}_{2..})$ were obtained by (a) averaging the values of $\gamma_{1_{ij}}$ and $\gamma_{2_{ij}}$ for the variate with the values of $\gamma_{1_{ij}}$ and $\gamma_{2_{ij}}$ for the covariate in the respective ij-th group and then (b) averaging the values from step (a) across the IJ groups. The $\overline{\rho}_{..}$, $\overline{\gamma}_{1..}$, and $\overline{\gamma}_{2..}$ were subsequently averaged across 10,000 (replications) × 21 (effect size) situations to obtain the overall averages; $\overline{\rho}$, $\overline{\gamma}_1$, and $\overline{\gamma}_2$ which appear in Tables II and III. The Headrick and Sawilowsky (1999) procedure produced excellent agreement between $\bar{\rho}$, $\bar{\gamma}_1$, and $\bar{\gamma}_2$ and the population parameters considered.

TABLE II

Average values of variate/covariate correlation ($\bar{\rho}$) computed in the simulation.

		Distribution				
		1	2	3	4	5
n	ρ	$\bar{\rho}$	$\bar{\rho}$	$\bar{\rho}$	$\bar{\rho}$	$\bar{\rho}$
3	.00	.000	-.001	.000	.002	-.003
	.30	.299	.300	.301	.299	.302
	.60	.600	.598	.599	.600	.601
	.90	.901	.901	.898	.899	.900
10	.00	.000	.000	-.001	.000	.002
	.30	.300	.300	.298	.301	.299
	.60	.600	.599	.601	.601	.602
	.90	.900	.899	.900	.900	.900
30	.00	.000	.000	-.001	.000	.001
	.30	.300	.299	.298	.300	.300
	.60	.600	.600	.600	.599	.601
	.90	.900	.900	.901	.900	.899

Note. n denotes the sample size. Values of $\bar{\rho}$ were based on 210,000 repetitions. The distributions are: (1) standard normal ($\mu = 0;\ \sigma = 1;\ \gamma_1 = 0;\ \gamma_2 = 0$);
(2) symmetric and light-tailed ($\mu = 0;\ \sigma = 1;\ \gamma_1 = 0;\ \gamma_2 = -\ 1.15132$);
(3) symmetric and extremely heavy-tailed ($\mu = 0;\ \sigma = 1;\ \gamma_1 = 0;\ \gamma_2 = 25$);
(4) asymmetric and moderately heavy-tailed ($\mu = 0;\ \sigma = 1;\ \gamma_1 = 1.633;\ \gamma_2 = 4$);
(5) extremely asymmetric and heavy-tailed ($\mu = 0;\ \sigma = 1;\ \gamma_1 = \sqrt{8};\ \gamma_2 = 12$).

The Type I error and power analyses for the tests of interaction and homogeneity of regression coefficients are compiled in Table IV through Table XV. The column entries from left to right denote (a) the treatment effect size "c"; (b) Fi and Fh represent the parametric ANCOVA tests of interaction and homogeneity of regression coefficients; and Fi(r) and Fh(r) represent the RT ANCOVA tests of interaction and homogeneity of regression coefficients; (c) the nominal alpha levels; and (d) the proportion of rejections for the tests under the various levels of variate and covariate correlation and other parameters considered.

TABLE III

Average values of skew ($\bar{\gamma}_1$) and kurtosis ($\bar{\gamma}_2$) computed in the simulation.

Symmetric Distributions

γ_1	0.000		0.000		0.000	
γ_2		0.000		-1.151		25.000
n	$\bar{\gamma}_1$	$\bar{\gamma}_2$	$\bar{\gamma}_1$	$\bar{\gamma}_2$	$\bar{\gamma}_1$	$\bar{\gamma}_2$
3	-0.002	0.003	0.001	-1.150	-0.002	25.002
10	0.001	0.000	-0.002	-1.152	-0.001	25.000
30	0.000	0.000	-0.001	-1.151	0.000	25.001

Asymmetric Distributions

γ_1	1.633		$\sqrt{8}$	
γ_2		4.000		12.000
n	$\bar{\gamma}_1$	$\bar{\gamma}_2$	$\bar{\gamma}_1$	$\bar{\gamma}_2$
3	1.635	4.002	2.830	12.002
10	1.632	4.000	2.826	12.001
30	1.633	4.001	2.827	12.001

Note. n denotes the sample size. The average values of skew ($\bar{\gamma}_1$) and kurtosis ($\bar{\gamma}_2$) were each based on 210,000 repetitions.

Type I Error: Normal Distribution. The Type I error rates are compiled in Table IV for $n = 30$. As expected, the parametric F tests maintained Type I error rates close to nominal alpha levels for the tests of interaction and homogeneity of regression slopes. They were within the closed interval of $\alpha \pm 1.96\sqrt{\alpha(1-\alpha)/10000}$.

However, the results in Table IV indicate that the RT produced extremely liberal Type I error rates for values of $c \geq 0.80$. For example, with $\rho = .90$ and nominal alpha equal .05, the Type I error rates reached as high as .998 and 1.00 for the RT tests of interaction and homogeneity of regression coefficients, respectively.

Type I Error: Nonnormal Distributions. As indicated by the entries in Tables V, VI, and VII, when the conditional distributions were nonnormal, both procedures displayed unreasonable and erratic Type I error rates for the test of homogeneity of regression coefficients. With respect to the F test, the Type I error rates were extremely liberal when the conditional distribution being simulated possessed positive γ_2 and ultra-conservative when the conditional distribution possessed negative γ_2. In general, the degree of inflation or conservatism for both procedures was contingent on the set of parameters being simulated.

With respect to the test for interaction, the results in Tables V, VI, and VII, indicate that the F test was robust with respect to Type I error. However, the RT test for interaction exhibited more severely inflated Type I error rates at smaller values of c than for the case of when the conditional distribution was normal. For example, in Table VII with the variate and covariate distributed with values of γ_1 = 0.00, γ_2 = 25, c = 0.30, and ρ = .90, the Type I error rate was .253 for the RT ANCOVA.

The other conditional nonnormal distribution modeled, but not presented in the tables, also resulted in similar Type I error rate inflations for the RT test for interaction. More generally, there was a pattern of Type I error inflation: *ceteris paribus*, the larger the departure from normality, the more severe the Type I error inflations became for the RT test.

With regard to the other treatment effects modeled, the RT ANCOVA was robust with respect to the test of interaction only under certain circumstances. This occurred when (a) all effects were null, (b) the treatment effect pattern contained only an interaction, or (c) if and only if one main effect was present.

Power: Normal Distribution. As anticipated, the results compiled in Tables VIII and IX indicate a comparative power advantage for the F test when the conditional distribution was normal. For small effect sizes and low variate/covariate correlations, the RT rejected at a rate slightly below the F test.

However, for moderate to large effect sizes and moderate to strong variate/covariate correlations, the RT exhibited a power loss. For example, the results in Table VIII indicate that for values of, c = 1.30, and ρ = .90 the F test was rejecting at a rate of .788, but the RT was rejecting at a rate of only .572. Similar power losses for the RT are reported in Table IX for n = 10.

TABLE IV

Type I Error Rates for Interaction and Homogeneity of Regression Coefficients for the model $\tau_3 = \alpha_1 = c$; and $\tau_1 = \tau_2 = \tau_4 = \alpha_3 = -c$. $Y = X$ = Standard Normal Distribution. Sample size is $n = 30$.

			Correlation (ρ)			
c	Statistic	α	0.0	0.3	0.6	0.9
0.30	Fi	.050	.050	.050	.053	.050
		.010	.010	.010	.011	.011
	Fi(r)	.050	.052	.050	.052	.052
		.010	.010	.009	.010	.010
	Fh	.050	.048	.048	.047	.049
		.010	.010	.010	.010	.010
	Fh(r)	.050	.049	.043	.025	.032
		.010	.009	.008	.004	.007
0.80	Fi	.050	.053	.049	.050	.050
		.010	.011	.010	.010	.011
	Fi(r)	.050	.116	.127	.167	.435
		.010	.030	.035	.050	.225
	Fh	.050	.050	.047	.051	.051
		.010	.010	.010	.010	.011
	Fh(r)	.050	.054	.066	.174	.973
		.010	.012	.016	.051	.951
1.30	Fi	.050	.051	.048	.052	.048
		.010	.010	.010	.010	.011
	Fi(r)	.050	.781	.826	.930	.998
		.010	.491	.555	.770	.991
	Fh	.050	.047	.048	.051	.052
		.010	.009	.011	.010	.012
	Fh(r)	.050	.053	.093	.466	1.00
		.010	.013	.025	.218	1.00

Note. Fi, Fh denote the parametric tests and Fi(r) and Fh(r) denote the rank transform (RT) tests for Interaction and Homogeneity of Regression Slopes, respectively. α denotes nominal alpha. 10,000 repetitions were used to generate the tabled values.

Power: Nonnormal Distributions. The power results for nonnormal distributions are reported in Tables X through XV. As indicated in Tables X and XI, where the conditional distribution was symmetric and light-tailed, the parametric F test held a power advantage over the RT. The power advantage in

TABLE V

Type I Error Rates for Interaction and Homogeneity of Regression Coefficients for the model $\tau_3 = \alpha_1 = c$; and $\tau_1 = \tau_2 = \tau_4 = \alpha_3 = -c$. $Y = X =$ ($\mu = 0; \sigma = 1; \gamma_1 = 0;$ $\gamma_2 = -1.15132$). Sample size is $n = 30$.

			Correlation (ρ)			
c	Statistic	α	0.0	0.3	0.6	0.9
0.30	Fi	.050	.052	.049	.053	.056
		.010	.012	.010	.012	.012
	Fi(r)	.050	.052	.051	.053	.061
		.010	.012	.010	.012	.013
	Fh	.050	.054	.046	.023	.009
		.010	.011	.009	.005	.004
	Fh(r)	.050	.051	.043	.019	.007
		.010	.011	.008	.002	.004
0.80	Fi	.050	.055	.055	.052	.052
		.010	.012	.013	.010	.012
	Fi(r)	.050	.096	.101	.119	.338
		.010	.025	.027	.034	.159
	Fh	.050	.054	.043	.025	.009
		.010	.013	.008	.006	.005
	Fh(r)	.050	.058	.063	.140	.994
		.010	.013	.013	.035	.964
1.30	Fi	.050	.053	.057	.054	.051
		.010	.011	.014	.012	.010
	Fi(r)	.050	.708	.763	.900	.998
		.010	.385	.462	.704	.990
	Fh	.050	.054	.044	.024	.009
		.010	.011	.009	.006	.004
	Fh(r)	.050	.064	.099	.489	1.00
		.010	.016	.026	.218	1.00

Note. Fi, Fh denote the parametric tests and Fi(r) and Fh(r) denote the rank transform (RT) tests for Interaction and Homogeneity of Regression Slopes, respectively. α denotes nominal alpha. 10,000 repetitions were used to generate the tabled values.

favor of the F test became more pronounced as either the effect size or strength of the variate/covariate correlation increased. However, the results reported for the extremely asymmetric/heavy-tailed and symmetric/extremely heavy-tailed distributions differed from the results of the standard normal and symmetric/light-

TABLE VI

Type I Error Rates for Interaction and Homogeneity of Regression Coefficients for the model $\tau_3 = \alpha_1 = c$; and $\tau_1 = \tau_2 = \tau_4 = \alpha_3 = -c$. Y=X=($\mu = 0; \sigma = 1$; $\gamma_1 = \sqrt{8}$; $\gamma_2 = 12$). Sample size is $n = 30$.

			Correlation (ρ)			
c	Statistic	α	0.0	0.3	0.6	0.9
0.30	Fi	.050	.052	.048	.051	.051
		.010	.009	.011	.012	.012
	Fi(r)	.050	.092	.091	.111	.196
		.010	.026	.025	.032	.074
	Fh	.050	.095	.565	.829	.926
		.010	.046	.403	.706	.847
	Fh(r)	.050	.069	.129	.403	1.00
		.010	.018	.040	.190	.998
0.80	Fi	.050	.049	.052	.051	.052
		.010	.010	.010	.010	.011
	Fi(r)	.050	.515	.546	.655	.881
		.010	.258	.282	.404	.739
	Fh	.050	.097	.561	.828	.923
		.010	.046	.390	.697	.847
	Fh(r)	.050	.068	.182	.627	1.00
		.010	.020	.068	.363	.999
1.30	Fi	.050	.051	.050	.050	.050
		.010	.010	.012	.009	.012
	Fi(r)	.050	.972	.981	.991	.998
		.010	.888	.911	.954	.993
	Fh	.050	.094	.561	.819	.919
		.010	.042	.390	.693	.838
	Fh(r)	.050	.070	.187	.659	1.00
		.010	.020	.070	.392	1.00

Note. Fi, Fh denote the parametric tests and Fi(r) and Fh(r) denote the rank transform (RT) tests for Interaction and Homogeneity of Regression Slopes, respectively. α denotes nominal alpha. 10,000 repetitions were used to generate the tabled values.

tailed distributions. Specifically, when the effect sizes modeled were weak, the RT ANCOVA held a decisive power advantage over the F test for these nonnormal distributions. For example, when the conditional distribution had values of $\gamma_1 = \sqrt{8}$, $\gamma_2 = 12$, $n = 3$, $c = 0.30$, and $\rho = .90$, the results in Table XII

TABLE VII

Type I Error Rates for Interaction and Homogeneity of Regression Coefficients for the model $\tau_3 = \alpha_1 = c$; and $\tau_1 = \tau_2 = \tau_4 = \alpha_3 = -c$. $Y = X = (\mu = 0; \sigma = 1; \gamma_1 = 0;$ $\gamma_2 = 25)$. Sample size is $n = 30$.

			Correlation (ρ)			
c	Statistic	α	0.0	0.3	0.6	0.9
0.30	F_i	.050	.052	.045	.052	.052
		.010	.011	.010	.011	.010
	$F_{i(r)}$	.050	.090	.091	.118	.253
		.010	.025	.022	.035	.118
	F_h	.050	.113	.471	.902	.985
		.010	.061	.326	.821	.967
	$F_{h(r)}$	.050	.055	.065	.151	.602
		.010	.012	.016	.051	.425
0.80	F_i	.050	.049	.049	.050	.052
		.010	.010	.010	.009	.010
	$F_{i(r)}$	.050	.620	.680	.822	.952
		.010	.361	.443	.639	.888
	F_h	.050	.113	.463	.900	.983
		.010	.061	.320	.820	.965
	$F_{h(r)}$	.050	.058	.131	.545	.944
		.010	.015	.040	.327	.885
1.30	F_i	.050	.047	.052	.048	.052
		.010	.008	.010	.009	.011
	$F_{i(r)}$	.050	.987	.990	.998	1.00
		.010	.937	.963	.990	.999
	F_h	.050	.112	.460	.899	.987
		.010	.061	.319	.820	.966
	$F_{h(r)}$	.050	.062	.179	.743	.982
		.010	.017	.063	.529	.961

Note. F_i, F_h denote the parametric tests and $F_{i(r)}$ and $F_{h(r)}$ denote the rank transform (RT) tests for Interaction and Homogeneity of Regression Slopes, respectively. α denotes nominal alpha. 10,000 repetitions were used to generate the tabled values.

TABLE VIII

Power Analysis for the test of Interaction and for the model $(\alpha\tau)_{21} = (\alpha\tau)_{24} = c$. $Y = X$ = Standard Normal Distribution. Sample size is $n = 3$.

			Correlation (ρ)			
c	Statistic	α	0.0	0.3	0.6	0.9
0.30	Fi	.050	.056	.056	.058	.076
		.010	.011	.011	.013	.017
	Fi(r)	.050	.0575	.056	.056	.070
		.010	.012	.013	.014	.015
0.80	Fi	.050	.091	.093	.118	.341
		.010	.022	.025	.030	.133
	Fi(r)	.050	.090	.093	.108	.250
		.010	.023	.023	.029	.036
1.30	Fi	.050	.177	.191	.268	.788
		.010	.053	.062	.100	.525
	Fi(r)	.050	.156	.170	.220	.572
		.010	.047	.054	.074	.305

Note. F_i and $F_{i(r)}$ denote the parametric and rank transform (RT) tests. α denotes nominal alpha. 10,000 repetitions were employed to generate the tabled values.

TABLE IX

Power Analysis for the test of Interaction and for the model $(\alpha\tau)_{21} = (\alpha\tau)_{24} = c$. $Y = X$ = Standard Normal Distribution. Sample size is $n = 10$.

			Correlation (ρ)			
c	Statistic	α	0.0	0.3	0.6	0.9
0.30	Fi	.050	.073	.075	.090	.202
		.010	.016	.016	.020	.069
	Fi(r)	.050	.073	.074	.086	.176
		.010	.016	.017	.020	.055
0.80	Fi	.050	.278	.305	.419	.950
		.010	.104	.116	.198	.839
	Fi(r)	.050	.250	.272	.377	.884
		.010	.097	.103	.167	.713
1.30	Fi	.050	.676	.722	.880	1.00
		.010	.418	.477	.706	1.00
	Fi(r)	.050	.612	.655	.812	.999
		.010	.347	.397	.595	.994

Note. F_i and $F_{i(r)}$ denote the parametric and rank transform (RT) tests. α denotes nominal alpha. 10,000 repetitions were employed to generate the tabled values.

TABLE X

Power Analysis for the test of Interaction and for the model $(\alpha\tau)_{21} = (\alpha\tau)_{24} = c$. $Y = X = (\mu = 0; \sigma = 1; \gamma_1 = 0; \gamma_2 = -1.15132)$. Sample size is $n = 3$.

			Correlation (ρ)			
c	Statistic	α	0.0	0.3	0.6	0.9
0.30	Fi	.050	.056	.057	.061	.076
		.010	.012	.012	.014	.016
	Fi(r)	.050	.058	.056	.059	.072
		.010	.012	.011	.013	.014
0.80	Fi	.050	.094	.098	.115	.337
		.010	.022	.025	.030	.135
	Fi(r)	.050	.090	.089	.104	.247
		.010	.020	.024	.028	.087
1.30	Fi	.050	.170	.172	.257	.775
		.010	.050	.052	.092	.517
	Fi(r)	.050	.140	.155	.204	.556
		.010	.038	.046	.066	.280

Note. Fi and Fi(r) denote the parametric and rank transform (RT) tests. α denotes nominal alpha. 10,000 repetitions were employed to generate the tabled values.

TABLE XI

Power Analysis for the test of Interaction and for the model $(\alpha\tau)_{21} = (\alpha\tau)_{24} = c$. $Y = X = (\mu = 0; \sigma = 1; \gamma_1 = 0; \gamma_2 = -1.15132)$. Sample size is $n = 10$.

			Correlation (ρ)			
c	Statistic	α	0.0	0.3	0.6	0.9
0.30	Fi	.050	.076	.078	.088	.201
		.010	.017	.016	.021	.068
	Fi(r)	.050	.076	.075	.083	.181
		.010	.017	.017	.020	.059
0.80	Fi	.050	.270	.292	.424	.943
		.010	.101	.112	.194	.826
	Fi(r)	.050	.226	.243	.352	.869
		.010	.079	.089	.146	.682
1.30	Fi	.050	.674	.733	.877	1.00
		.010	.418	.469	.707	.999
	Fi(r)	.050	.549	.593	.769	.999
		.010	.294	.335	.541	.995

Note. Fi and Fi(r) denote the parametric and rank transform (RT) tests. α denotes nominal alpha. 10,000 repetitions were employed to generate the tabled values.

TABLE XII

Power Analysis for the test of Interaction and for the model $(\alpha\tau)_{21} = (\alpha\tau)_{24} = c$. $Y = X = (\mu = 0; \sigma = 1;\ \gamma_1 = \sqrt{8}; \gamma_2 = 12)$. Sample size is $n = 3$.

			Correlation (ρ)			
c	Statistic	α	0.0	0.3	0.6	0.9
0.30	Fi	.050	.050	.051	.051	.087
		.010	.013	.012	.014	.020
	Fi(r)	.050	.084	.086	.090	.140
		.010	.021	.021	.022	.039
0.80	Fi	.050	.094	.104	.139	.428
		.010	.025	.027	.040	.206
	Fi(r)	.050	.147	.157	.183	.400
		.010	.043	.046	.057	.161
1.30	Fi	.050	.230	.247	.340	.824
		.010	.080	.092	.146	.619
	Fi(r)	.050	.221	.245	.291	.601
		.010	.075	.091	.104	.318

Note. Fi and Fi(r) denote the parametric and rank transform (RT) tests. α denotes nominal alpha. 10,000 repetitions were employed to generate the tabled values.

TABLE XIII

Power Analysis for the test of Interaction and for the model $(\alpha\tau)_{21} = (\alpha\tau)_{24} = c$. $Y = X = (\mu = 0; \sigma = 1;\ \gamma_1 = \sqrt{8}; \gamma_2 = 12)$. Sample size is $n = 10$.

			Correlation (ρ)			
c	Statistic	α	0.0	0.3	0.6	0.9
0.30	Fi	.050	.072	.072	.086	.214
		.010	.016	.015	.020	.075
	Fi(r)	.050	.175	.187	.226	.552
		.010	.056	.060	.076	.295
0.80	Fi	.050	.287	.331	.457	.945
		.010	.113	.142	.233	.849
	Fi(r)	.050	.582	.612	.735	.987
		.010	.314	.347	.469	.943
1.30	Fi	.050	.692	.742	.878	1.00
		.010	.458	.527	.724	.999
	Fi(r)	.050	.845	.876	.941	1.00
		.010	.625	.670	.803	.997

Note. Fi and Fi(r) denote the parametric and rank transform (RT) tests. α denotes nominal alpha. 10,000 repetitions were employed to generate the tabled values.

TABLE XIV

Power Analysis for the test of Interaction and for the model $(\alpha\tau)_{21} = (\alpha\tau)_{24} = c$. $Y = X = (\mu = 0; \sigma = 1; \gamma_1 = 0; \gamma_2 = 25)$. Sample size is $n = 3$.

			Correlation (ρ)			
c	Statistic	α	0.0	0.3	0.6	0.9
0.30	Fi	.050	.038	.044	.052	.097
		.010	.007	.008	.012	.029
	Fi(r)	.050	.076	.080	.103	.212
		.010	.019	.021	.028	.071
0.80	Fi	.050	.104	.112	.162	.517
		.010	.025	.030	.049	.280
	Fi(r)	.050	.167	.191	.283	.637
		.010	.055	.062	.104	.390
1.30	Fi	.050	.251	.277	.406	.854
		.010	.093	.111	.194	.703
	Fi(r)	.050	.252	.291	.436	.518
		.010	.090	.109	.201	.606

Note. Fi and Fi(r) denote the parametric and rank transform (RT) tests. α denotes nominal alpha. 10,000 repetitions were employed to generate the tabled values.

TABLE XV

Power Analysis for the test of Interaction and for the model $(\alpha\tau)_{21} = (\alpha\tau)_{24} = c$. $Y = X = (\mu = 0; \sigma = 1; \gamma_1 = 0; \gamma_2 = 25)$. Sample size is $n = 10$.

			Correlation (ρ)			
c	Statistic	α	0.0	0.3	0.6	0.9
0.30	Fi	.050	.069	.074	.090	.250
		.010	.015	.016	.020	.090
	Fi(r)	.050	.178	.204	.306	.821
		.010	.061	.068	.124	.613
0.80	Fi	.050	.313	.346	.491	.945
		.010	.130	.152	.264	.860
	Fi(r)	.050	.629	.703	.888	.999
		.010	.382	.461	.719	.997
1.30	Fi	.050	.707	.753	.882	.999
		.010	.493	.548	.747	.997
	Fi(r)	.050	.877	.915	.986	1.00
		.010	.694	.772	.942	.999

Note. Fi and Fi(r) denote the parametric and rank transform (RT) tests. α denotes nominal alpha. 10,000 repetitions were employed to generate the tabled values.

indicate that the proportion of rejections for the RT ANCOVA was .140, whereas the proportion of rejections for the F test was .087. Similarly, when the conditional distribution had values of $\gamma_1 = 0$, $\gamma_2 = 25$, $n = 10$, $c = 0.30$, and $\rho = .60$, the results in Table XV indicate that the proportion of rejections for the RT ANCOVA was .306, whereas the proportion of rejections for the F test was .090.

However, the RT ANCOVA loses its power advantage over the F test as either the effect size and/or correlation increase. For example, when the conditional distribution had values of $\gamma_1 = 0$, $\gamma_2 = 25$, $n = 3$, $c = 1.30$, and $\rho = .90$, the results in Table XIV indicate that the proportion of rejections for the RT was only .518, whereas the proportion of rejections for the F test was .854.

The pattern that was pointed out with respect to the RT's Type I error inflations for the test of interaction was also evident in terms of power. That is, the larger the departure was from conditional normality, the more severe the power loss for the RT. Similarly, when the RT was robust with respect to Type I error, the RT maintained a power advantage over the parametric F test when either (a) the treatment effect pattern contained only an interaction, or (b) if and only if one main effect was present.

5. DISCUSSION

With respect to the test of homogeneity of regression coefficients, violating the assumption of conditional normality adversely affects the parametric ANCOVA's Type I error rates. Contingent on the conditional distribution and strength of variate/covariate correlation, the lack of robustness of the F test ranged from ultra-conservative to ultra-liberal. Thus, the results of this study invalidate the parametric F test as a test for parallelism when the variate <u>and</u> covariate are nonnormally distributed. Conover and Iman (1982, Table 4) also reported increased Type I error rates for this test when the variate followed either a lognormal or Cauchy distribution. For example, when the distribution was

Cauchy, Conover and Iman (1982) reported a Type I error rate of .103 ($\alpha = 0.05$). It should be noted that the covariate was normally distributed for all experimental situations in the Conover and Iman (1982) study.

Although the focus in this study was on moderate to severe departures from conditional normality, Headrick (1997) also demonstrated empirically that even for small departures from conditional normality, the F test produced either conservative or liberal Type I error rates for the test of equal slopes. For example, when the variate and covariate possessed values of $\gamma_1 = 0.0$, $\gamma_2 = -0.34$ or $\gamma_1 = 0.0$, $\gamma_2 = 0.31$, with $\rho = .60$, and for large sample sizes, the Type I error rates reported were .020 and .100.

Unlike the parametric F test for parallelism, the RT test failed when both the variate and covariate were normally distributed. Further, the erratic Type I error rates associated with the RT depended not only on the conditional distribution and strength of the variate/covariate correlation, but also the treatment effect pattern being modeled. Furthermore, in the absence of any treatment effect, the RT became ultra-conservative as the strength of the variate/covariate correlation and/or sample size increased. Consequently, such conservative Type I error rates leave the researcher conducting the RT ANCOVA test without guidance from a preliminary test of underlying assumptions. Thus, and like the parametric test, the results of this study also invalidate the RT as a test for the homogeneity of regression coefficients.

With respect to the test of interaction, the RT procedure resulted in severe Type I error inflations. As a result, it is interesting to compare the RT and parametric procedures in terms of their expected group means to show how the parametric procedure is not invariant under monotone transformations. In terms of this experiment, the nonlinear nature of the RT reversed the absence of interaction in the original scores when both main effects were present.

In order to demonstrate this, it is only necessary to consider the case where the *IJ* groups in (1) have an absolutely continuous normal distribution. The stochastic disturbance terms in (1) are assumed to be independently and identically distributed with zero means and unit variances.

Specifically, if y denotes an observation of the Y_{ijk} from the *ij*-th group in (1) and $\lambda(i,j,k) = 1$ or 0 indicating the scores of the $Y_{ijk} < y$ or $Y_{ijk} > y$, then the rank of y (R_y) can be defined as,

$$R_y = 1 + \sum_i \sum_j \sum_k \lambda(i,j,k). \tag{6}$$

It follows that the expected value of R_y in the *ij*-th group can be expressed as,

$$E[R_y] = 1 + \sum_i \sum_j \sum_k \Pr\{Y_{ijk} < y\}. \tag{7}$$

Suppose y is any observation of the Y_{ijk} in the $\alpha\tau$-th group from (1). Then (7) can be rewritten to express the expected value of R_y as,

$$E[R_y] = 1 + \frac{1}{2}(n-1) + \sum_{i \neq \alpha} \sum_{j \neq \tau} \sum_k \Pr\{Y_{ijk} < y\}, \tag{8}$$

where the value $\frac{1}{2}(n-1)$ indicates the sum of $\Pr\{Y_{ijk} < y\}$ for the $Y_{ijk} \neq y$ in the $\alpha\tau$-th group.

Because the Y_{ijk} have the same expectation within their respective *ij*-th group, (8) can be simplified by expressing the term $\sum_{i \neq \alpha} \sum_{j \neq \tau} \sum_k \Pr\{Y_{ijk} < y\}$ in terms of the population means for the Y_{ijk}. Thus, let $\mu_{\alpha\tau}$ denote the population mean of the Y_{ijk} in the $\alpha\tau$-th group. Further, let μ_{ij} denote the population means of the other $IJ - 1$ groups of the Y_{ijk} not belonging to the $\alpha\tau$-th group. It follows that (8) can be expressed in terms of the expected group means for the ranks as

$$E[R_{\alpha\tau}] = 1 + \frac{1}{2}(n-1) + n \sum_{i \neq \alpha} \sum_{j \neq \tau} \Pr\{\mu_{ij} < \mu_{\alpha\tau}\}. \tag{9}$$

To determine the $\sum_{i\neq\alpha}\sum_{j\neq\tau}\Pr\{\mu_{ij} < \mu_{\alpha\tau}\}$ in (9) let

$$z_{ij} = \frac{\mu_{\alpha\tau} - \mu_{ij}}{\sqrt{2}}, \qquad \forall_{ij\neq\alpha\tau}. \tag{10}$$

Thus, the expected value of the ranks in the $\alpha\tau$ -th group can be computed as,

$$E[R_{\alpha\tau}] = 1 + \frac{1}{2}(n-1) + n\sum_i\sum_j\int_{-\infty}^{z_{ij}}\frac{1}{\sqrt{2\pi}}e^{\frac{-w^2}{2}}dw. \tag{11}$$

A comparison of the expected group means and their corresponding values of interaction between the original scores and their associated ranks are presented in Table XVI through Table XIX for the case of n = 30, $\tau_3 = \alpha_1 = 0.80$, and $\tau_1 = \tau_2 = \tau_3 = \tau_4 = -\ 0.80$. This model indicates both main effects are nonnull and the interaction is null. This example is also associated with the results presented in Table IV. By inspection of Table XVII and Table XIX, note that the original scores indicate no interaction while their ranks indicate an interaction.

To illustrate the computation of an expected group mean for the ranked data, consider the group $\alpha_1\tau_3$ in Table XVIII where the expected group mean is 316.352. This value was determined by the following steps: (a) use Equation (10) to compute the values of the z_{ij},

these values are: $z_{11} = \frac{1.60 - 0.00}{\sqrt{2}}$; $z_{12} = \frac{1.60 - 0.00}{\sqrt{2}}$; $z_{14} = \frac{1.60 - 0.00}{\sqrt{2}}$;

$z_{21} = \frac{1.60 - (-\ 0.80)}{\sqrt{2}}$; $z_{22} = \frac{1.60 - (-\ 0.80)}{\sqrt{2}}$; $z_{23} = \frac{1.60 - 0.80}{\sqrt{2}}$; $z_{24} = \frac{1.60 - (-\ 0.80)}{\sqrt{2}}$;

$z_{31} = \frac{1.60 - (-\ 1.60)}{\sqrt{2}}$; $z_{32} = \frac{1.60 - (-\ 1.60)}{\sqrt{2}}$; $z_{33} = \frac{1.60 - 0.00}{\sqrt{2}}$; $z_{34} = \frac{1.60 - (-\ 1.60)}{\sqrt{2}}$;

and (b), compute the expected group mean for the ranked data by entering each value from step (a) into the upper limit of the integral in Equation (11). Hence,

$$E[R_{13}] = 316.352 = 1 + \frac{1}{2}(30-1) + 30\begin{Bmatrix} .8710 + .8710 + .8710 + .9552 + \\ .9552 + .7142 + .9552 + .9882 + \\ .9882 + .8710 + .9882 \end{Bmatrix}.$$

TABLE XVI

Expected Group Means of the raw scores; $\tau_3 = \alpha_1 = c$ and $\tau_1 = \tau_2 = \tau_4 = \alpha_3 = -c$; $c = 0.80$; $Y = X =$ Standard Normal Distribution; and $n = 30$ as in Table IV.

	τ_1	τ_2	τ_3	τ_4
α_1	0.000	0.000	1.600	0.000
α_2	-0.800	-0.800	0.800	-0.800
α_3	-1.600	-1.600	0.000	-1.600

TABLE XVII

Interactions of the raw scores; $\tau_3 = \alpha_1 = c$ and $\tau_1 = \tau_2 = \tau_4 = \alpha_3 = -c$; $c = 0.80$; $Y = X =$ Standard Normal Distribution; and $n = 30$ as in Table IV.

	τ_1	τ_2	τ_3	τ_4
α_1	0.000	0.000	0.000	0.000
α_2	0.000	0.000	0.000	0.000
α_3	0.000	0.000	0.000	0.000

TABLE XVIII

Expected Group Means of the ranks; $\tau_3 = \alpha_1 = c$ and $\tau_1 = \tau_2 = \tau_4 = \alpha_3 = -c$; $c = 0.80$; $Y = X =$ Standard Normal Distribution; and $n = 30$ as in Table IV.

	τ_1	τ_2	τ_3	τ_4
α_1	215.615	215.615	316.352	215.615
α_2	149.288	149.288	274.136	149.288
α_3	88.396	88.396	215.615	88.396

TABLE XIX

Interactions of the ranks; $\tau_3 = \alpha_1 = c$ and $\tau_1 = \tau_2 = \tau_4 = \alpha_3 = -c$; $c = 0.80$; $Y = X =$ Standard Normal Distribution; and $n = 30$ as in Table IV.

	τ_1	τ_2	τ_3	τ_4
α_1	4.216	4.216	-12.648	4.216
α_2	-1.812	-1.812	5.436	-1.812
α_3	-2.404	-2.404	7.213	-2.404

The derivation of the expected group means of the ranked scores demonstrates how the F test is not invariant under monotone transformations in the sense that the probability structure of the original problem of testing for no interaction is changed. Similar points were made by Blair, Sawilowsky, and Higgins (1987) and Thompson (1991) with respect to the RT tests of interaction for ANOVA designs. However, Type I error inflations can be much more pronounced for smaller effect sizes or for smaller sample sizes when the inclusion of a covariate is introduced. The results of this study invalidate the use of the RT ANCOVA as an alternative to the F test of interaction in factorial ANCOVA.

6. CONCLUSION

Researchers and publishers of statistical software continue to recommend the rank transformation (RT) as an alternative to parametric analysis. For example, Kleinbaum, Kupper, Muller, and Nizam (1998) stated, "If the problem concerns nonnormal distributions, methods of rank analysis (Conover and Iman 1981)...may be appropriate" (p. 249).

Similarly, the SAS (1995) and IMSL (1994) statistical packages promote the use of the RT procedures with respect to general linear models without restriction. For example, the most recent SAS manual states, "Many nonparametric statistical methods use ranks rather than original values of a variable. For example, a set of data can be passed through PROC RANK to obtain the ranks for a response variable that could then be fit to an analysis-of-variance model using the ANOVA or GLM procedures" (p. 493). The IMSL (1994) manual states, "Many of the tests described in this chapter may be computed using the routines described in other chapters after first substituting ranks for the observed values" (p. 582).

Based on results of this investigation, it is recommended that the rank transformation procedure not be used as either a test for interaction or a test for homogeneity of regression slopes in factorial ANCOVA. This is based on the severely inflated Type I error rates when the null hypothesis of interaction was true and while both main effects were nonnull - regardless of the conditional distribution or sample size being simulated. Similarly, the stronger the correlation between the variate and covariate, the more serious was the Type I error inflation. Further research on the comparative properties of other nonparametric tests, such as Hettmansperger (1984), Puri & Sen (1969a), and the adjusted RT (Blair & Sawilowsky, 1990; Salter & Fawcett, 1993) is warranted to find alternatives to the parametric factorial ANCOVA.

BIBLIOGRAPHY

Akritas, M. G. (1990). The rank transform method in some two factor designs. *Journal of the American Statistical Association*, 85, 73-78.

Angermeier, P. L., & Winston, M. R. (1998), Local vs. regional influences on local diversity in stream fish communities of Virginia. *Ecology*, 79, 911-927.

Augner, M., Provenza, F. D., & Villalba, J. J. (1998). A rule of thumb in mammalian herbivores? *Animal Behaviour*, 56, 337-345.

Blair, R. C. (1987). *RANGEN*. Boca Raton, FL: IBM.

Blair, R. C., & Sawilowsky, S. S. (1990, April). *A test for interaction based on the rank transform*. Paper presented at the annual meeting of the American Educational Research Association, Boston.

Blair, R. C., Sawilowsky, S. S., & Higgins, J. J. (1987). Limitations of the rank transform in tests of interactions. *Communications in Statistics: Simulation and Computation*, 16, 1133-1145.

Brunner, E., & Neumann, N. (1986). Rank tests in 2 × 2 designs. *Statistica Neulandica*, 40, 251-272.

Choi, Y. H., (1998). A study of the power of the rank transform test in a 2^3 factorial experiment. *Communications in Statistics*, 27, 259-274.

Conover, W. J., & Iman, R. L. (1976). On some alternative procedures using ranks for the analysis of experimental designs. *Communications in Statistics*, 5, 1348-1368.

Conover, W. J., & Iman, R. L. (1981). Rank transformations as a bridge between parametric and nonparametric statistics. *American Statistician*, 35, 124-133.

Conover, W. J., & Iman, R. L. (1982). Analysis of covariance using the rank transformation. *Biometrics*, 38, 715-724.

Deshon, R. P., & Alexander, R. A. (1996). Alternative procedures for testing regression slope homogeneity when group error variances are unequal. *Psychological Methods*, 1, 261-277.

Fleishman, A. I. (1978). A method for simulating non-normal distributions. *Psychometrika*, 43, 521-532.

Harwell, M. R., & Serlin, R. C. (1988). An experimental study of a proposed test of non-parametric analysis of covariance. *Psychological Bulletin*, 104, 268-281.

Headrick, T. C. (1997). Type I error and power of the rank transform analysis of covariance (ANCOVA) in a 3 × 4 factorial layout. Unpublished doctoral dissertation, Wayne State University, Detroit, MI.

Headrick, T. C. & Sawilowsky, S. S. (1999). Simulating correlated multivariate nonnormal distributions: Extending the Fleishman power method. *Psychometrika*, 64, 25-35.

Hettmansperger, T. P. (1984). *Statistical inference based on ranks*. NY: Wiley.

Iman, R. L. (1974). A power study of a rank transform for the two way classification model when interaction may be present. *Canadian Journal of Statistics*, 2, 227-239.

Iman, R. L., & Conover, W. J. (1979). The use of rank transform in regression. *Technometrics*, 21, 499-509.

International Mathematical and Statistical Libraries. (1994). *IMSL library, reference manual*. Houston, TX: Author.

Kleinbaum, D. G., Kupper, L. L., Muller, K. E., & Nizam, A. (1998). *Applied regression analysis and other multivariable methods* (3rd). Pacific Grove, CA: Duxbury Press.

Knapp, T. R. & Sowyer, V. H. (1967). Some empirical results concerning the power of Bartlett's test of the significance of a correlation matrix. *American Educational Research Journal*, 4, 13-17.

Lahey, Computer Systems, Inc. (1994). *Personal Fortran (version 3.0)*. Incline Village, NV: Author.

Olejnik, S. F., & Algina, J. (1984). Parametric ANCOVA and the rank transform ANCOVA when the data are conditionally non-normal and heteroscedastic. *Journal of Educational Statistics*, 9, 129-150.

Olejnik, S. F., & Algina, J. (1985). A review of non-parametric alternatives to analysis of covariance. *Evaluation Review*, 9, 51-83.

Olejnik, S. F., & Algina, J. (1987). An analysis of statistical power of parametric ANCOVA and rank transform ANCOVA. *Communications in Statistics: Theory and Methods*, 16, 1923-1949.

Potvin, C., & Roff, D. A. (1993). Distribution-free and robust statistical methods: viable alternatives to parametric statistics. *Ecology*, 74, 1617-1628.

Puri, M. L., & Sen, P. K. (1969a). Analysis of covariance based on general rank scores. *Annals of Mathematical Statistics*, 40, 610-618.

Quade, D. (1967). Rank analysis of covariance. *Journal of the American Statistical Association*, 62, 1187-1200.

Regeth, R. A., & Stine, W. W. (1998). Comparing means from nonnormal distributions: The bisquare-weighted analysis of variance. *Behavior Research Methods, Instruments*, & Computers, 30, 707-712.

Salter, K. C., & Fawcett, R. F. (1993). The ART test of interaction: A robust and powerful rank test of interaction in factorial models. *Communications in Statistics: Computation and Simulation*, 22, 137-153.

SAS Institute, Inc. (1995). *SAS User's Guide: Statistics*, 5th ed. N.C.: SAS Institute.

Sawilowsky, S. S., Blair, R. C., & Higgins, J. J. (1989). An investigation of the Type I error and power properties of the rank transform procedure in factorial ANOVA. *Journal of Educational Statistics*, 14, 255-267.

Seaman, S., Algina, J., & Olejnik, S. F. (1985). Type I error probabilities and power of the rank and parametric ANCOVA procedures. *Journal of Educational Statistics*, 10, 345-367.

Stephenson, W. R., & Jacobson, D. (1988). A comparison of non-parametric analysis of covariance techniques. *Communication in Statistics: Computation and Simulation*, 26, 605-618.

Thompson, G. L. (1991). A note on the rank transform for interactions. *Biometrika*, 78, 697-701.

Visual Numerics. (1994). *IMSL Math/Library: FORTRAN subroutines for mathematical applications*, Volume II. Houston, TX: Author.

Winer, B. J., Brown, D. R., & Michels, K. M. (1991). *Statistical principles in experimental designs* (3rd). NY: McGraw-Hill.

AERA

U.S. Department of Education
Office of Educational Research and Improvement (OERI)
National Library of Education (NLE)
Educational Resources Information Center (ERIC)

ERIC®

TM030794

REPRODUCTION RELEASE

(Specific Document)

I. DOCUMENT IDENTIFICATION:

Title: Properties of the Rank Transformation in Factorial Analysis of Covariance

Author(s): TODD C. HEADRICK; SHLOMO S. SAWILOWSKY

Corporate Source: ~~[illegible]~~ Southern Illinois Univ, Carbondale

Publication Date: Post: 4-28-00

II. REPRODUCTION RELEASE:

In order to disseminate as widely as possible timely and significant materials of interest to the educational community, documents announced in the monthly abstract journal of the ERIC system, *Resources in Education* (RIE), are usually made available to users in microfiche, reproduced paper copy, and electronic media, and sold through the ERIC Document Reproduction Service (EDRS). Credit is given to the source of each document, and, if reproduction release is granted, one of the following notices is affixed to the document.

If permission is granted to reproduce and disseminate the identified document, please CHECK ONE of the following three options and sign at the bottom of the page.

The sample sticker shown below will be affixed to all Level 1 documents	The sample sticker shown below will be affixed to all Level 2A documents	The sample sticker shown below will be affixed to all Level 2B documents
PERMISSION TO REPRODUCE AND DISSEMINATE THIS MATERIAL HAS BEEN GRANTED BY Sample TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC) 1	PERMISSION TO REPRODUCE AND DISSEMINATE THIS MATERIAL IN MICROFICHE, AND IN ELECTRONIC MEDIA FOR ERIC COLLECTION SUBSCRIBERS ONLY, HAS BEEN GRANTED BY Sample TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC) 2A	PERMISSION TO REPRODUCE AND DISSEMINATE THIS MATERIAL IN MICROFICHE ONLY HAS BEEN GRANTED BY Sample TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC) 2B
Level 1 [X]	Level 2A [ ]	Level 2B [ ]
Check here for Level 1 release, permitting reproduction and dissemination in microfiche or other ERIC archival media (e.g., electronic) and paper copy.	Check here for Level 2A release, permitting reproduction and dissemination in microfiche and in electronic media for ERIC archival collection subscribers only	Check here for Level 2B release, permitting reproduction and dissemination in microfiche only

Documents will be processed as indicated provided reproduction quality permits.
If permission to reproduce is granted, but no box is checked, documents will be processed at Level 1.

I hereby grant to the Educational Resources Information Center (ERIC) nonexclusive permission to reproduce and disseminate this document as indicated above. Reproduction from the ERIC microfiche or electronic media by persons other than ERIC employees and its system contractors requires permission from the copyright holder. Exception is made for non-profit reproduction by libraries and other service agencies to satisfy information needs of educators in response to discrete inquiries.

Sign here, → please

Signature: Todd C. Headrick

Printed Name/Position/Title: TODD C. HEADRICK / ASSISTANT PROFESSOR

Organization/Address: Southern Illinois University at Carbondale MAILCODE 4618, CARBONDALE, IL 62901-4618

Telephone: 618-453-1822 FAX:

E-Mail Address: headrick@siu.edu Date: 4-18-00

(over)

ERIC® Clearinghouse on Assessment and Evaluation

University of Maryland
1129 Shriver Laboratory
College Park, MD 20742-5701

Tel: (800) 464-3742
(301) 405-7449
FAX: (301) 405-8134
ericae@ericae.net
http://ericae.net

March 2000

Dear AERA Presenter,

Congratulations on being a presenter at AERA. The ERIC Clearinghouse on Assessment and Evaluation would like you to contribute to ERIC by providing us with a written copy of your presentation. Submitting your paper to ERIC ensures a wider audience by making it available to members of the education community who could not attend your session or this year's conference.

Abstracts of papers accepted by ERIC appear in *Resources in Education (RIE)* and are announced to over 5,000 organizations. The inclusion of your work makes it readily available to other researchers, provides a permanent archive, and enhances the quality of *RIE*. Abstracts of your contribution will be accessible through the printed, electronic, and internet versions of *RIE*. The paper will be available **full-text, on demand through the ERIC Document Reproduction Service** and through the microfiche collections housed at libraries around the world.

We are gathering all the papers from the AERA Conference. We will route your paper to the appropriate clearinghouse and you will be notified if your paper meets ERIC's criteria. Documents are reviewed for contribution to education, timeliness, relevance, methodology, effectiveness of presentation, and reproduction quality. You can track our processing of your paper at **http://ericae.net.**

To disseminate your work through ERIC, you need to sign the reproduction release form on the back of this letter and include it with **two** copies of your paper. You can drop of the copies of your paper and reproduction release form at the ERIC booth (223) or mail to our attention at the address below. **If you have not submitted your 1999 Conference paper please send today or drop it off at the booth with a Reproduction Release Form.** Please feel free to copy the form for future or additional submissions.

Mail to: AERA 2000/ERIC Acquisitions
The University of Maryland
1129 Shriver Lab
College Park, MD 20742

Sincerely,

Lawrence M. Rudner

Lawrence M. Rudner, Ph.D.
Director, ERIC/AE

ERIC/AE is a project of the Department of Measurement, Statistics and Evaluation at the College of Education, University of Maryland.